METAGENOMIC ANALYSIS OF SUBSURFACE ARCHAEA OF THE MISCELLANEOUS CRENARCHAEOTAL GROUP

Andrew Hyde

A thesis submitted to the faculty at the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for graduating with Honors in the Department of Biology in the College of Arts and Sciences

Chapel Hill
2014

ABSTRACT

The subsurface biosphere may account for one-third to one-half of the total biomass on Earth; however, we have only recently begun to explore this frontier. The life found in the subsurface is largely bacterial and archaeal; and most of these organisms remain unstudied. Among phylotypes found in marine sediments, the Miscellaneous Crenarchaeotal Group (MCG) of archaea have emerged as dominant players. Despite its apparent ubiquity, hardly anything is known about this very diverse group of microbes. In this work, I have analyzed several complete archaeal genomes belonging to the MCG group from the subsurface of the White Oak River estuary. I found genes indicating that the MCG subgroups analyzed here have the ability to reduce nitrite and other oxidized nitrogen compounds. These results suggest that our understanding of marine biogeochemical cycling may need to be modified to account for the significant role these organisms may play.

INTRODUCTION

Nearly every place we have looked for life on Earth, we have found it. Persisting and thriving in the most unexpected places, life has been found in radioactive waste, ice cores from the arctic, and hydrothermal vents at the mid-Atlantic ridge. We have even found life well over 1 km below the surface of the Earth (Dong et al., 2014) . Subsurface[1]

[1] Terminology dealing with the subsurface is ambiguous and no clear distinctions exist. "Subsurface" is usually meant to imply that there is no direct contact with the water column or atmosphere. (Teske & Sørensen, 2007)

prokaryotes may account for between 1/10 to 1/3 of the Earth's total biomass (Whitman, Coleman, & Wiebe, 1998). Despite this astonishing fact, we have only recently begun to explore this new frontier in environmental microbiology.

Life buried below the surface of the Earth must deal with very interesting conditions and challenges to survival. With no light, organisms here must fix carbon by non-photosynthetic means, or assimilate exogenous organics, which typically increase in abundance with depth. In typical aquatic sediments, oxygen is depleted within the first few mm, meaning that all respiration below the surface must use an alternate terminal electron acceptor (TEA). Despite knowing this, there are still many unanswered questions: What are the metabolic capabilities of these organisms? What enzymes do they use? Though these are basic questions, we are still discovering new metabolic pathways to this day (e.g. anammox) (Strous et al., 1999).

Even though members of *Eukarya* (e.g. fungi and protists) have been found in the subsurface, life here is largely of the domains *Archaea* and *Bacteria* (Edgcomb, Beaudoin, Gast, Biddle, & Teske, 2011). While we have known of the existence of the *Bacteria* for well over a century, and the *Eukarya* for millennia, the discovery this third domain of life, the *Archaea*, is a comparatively recent event (Woese & Fox, 1977; Woese, Kandler, & Wheelis, 1990). Discovered in the late 20th century, these organisms have been found in a wide variety of habitats, from the extreme environments in which they were initially thought to exclusively reside, to within the gut and in nearly every other mesophilic environment (Miller, Wolin, Conway de Macario, & Macario, 1982; Delong, 1994)Both lipid and genetic analyses have shown that archaea contribute to a

significant proportion of the biomass in marine and subsurface sediments (Kubo et al., 2012; Lipp, Morono, Inagaki, & Hinrichs, 2008)

There are thought to be at least five phyla of *Archaea*: the *Euryarchaeota*, *Crenarchaeota*, *Nanoarchaeota*, *Thaumarchaeota*, and *Korarchaeota*. Of these, the *Crenarchaeota* and *Euryarchaeota* are the best characterized. Within the Crenarchaeota, there are many divisions; one of the most phylogenetically diverse groups is the Miscellaneous Crenarchaeotal Group (MCG). Within the MCG, there are over a dozen subdivisions[2], with 16S rRNA sequence similarities of only 76% from MCG-1 to MCG-17; the MCG may therefore be as diverse as all of the Proteobacteria (Kubo et al., 2012; Lazar, personal communication). MCG archaea have been found in a multitude of environments, from deep subsurface sediments to brackish waters, sulfidic springs to hot springs (Parkes et al., 2005; Hershberger, Barns, Reysenbach, Dawson, & Pace, 1996; Elshahed et al., 2003; Barns, Delwiche, Palmer, & Pace, 1996).

Despite their seeming ubiquity, not much is known about the metabolic capabilities of the MCG. The current working hypothesis is that the MCG are anaerobic heterotrophs (Biddle et al., 2006; Teske & Sørensen, 2007). Understanding the metabolic capabilities of such a predominant organism is important to understanding how the MCG might play a role in global biogeochemical cycles. This thesis aims to address the anaerobic metabolic capabilities of some of the MCG subdivisions.

The White Oak River Estuary (WOR), located off the coast of North Carolina, has an abundance of uncultured archaea, with many studies finding that MCG sequences dominate the samples (Gagen, Huber, Meador, Hinrichs, & Thomm, 2013; Kubo et al.,

[2] This is according to the categorization proposed by Kubo et al, 2012, Sørensen & Teske, 2006, and Nercessian, Bienvenu, Moreira, Prieur, & Jeanthon, 2005. Others have proposed other alternative subdivisions. See Jiang et al 2011, where a 6 subdivision categorization is proposed.

2012; K. G. Lloyd, Biddle, & Teske, 2011; Karen G. Lloyd, Alperin, & Teske, 2011). I have analyzed the genomes of MCG archaea from three sample sites in the WOR. At all three sites, the MCG were shown to be the predominant archaeal phylotype (Supplementary Figure 2). Here, I address the dissimilatory nitrogen and sulfur metabolic genes found in these genomes to try and determine the MCG metabolic capabilities with respect to inorganic compounds. Understanding the processes this significant portion of the biosphere are involved in will help us better understand how subsurface microbial communities impact global processes and how these could be exploited to potentially mitigate some of the anthropogenic damage done to the environment.

METHODS

A total of six, 1 meter core samples were collected from the White Oak River Estuary in North Carolina in October of 2010. There were three sampling sites (2 samples per site), (Figure 1). Each core was subsampled based on centimeters below sea floor (cmbsf): shallow (8-12 cm cmbsf), SMTZ (24-32 cmbsf), deep (44-54 cmbsf). The SMTZ (sulfate to methane transition zone) varies between habitats and was determined by geochemical measurements (Supplementary Figure 1).

Figure 1. Location of WOR estuary on the coast of North Carolina.

DNA extractions and high throughput Illumina sequencing for each subsample (shallow, SMTZ, deep) were done by Dr. Cassandre Lazar (Bennett, 2004). Dr. Brett

Baker processed the sequence data and assembled contigs and scaffolds from the sequence reads. The sequences were binned and then compared (Altschul, Gish, Miller, Myers, & Lipman, 1990) against three different databases using BLAST (Altschul et al., 1990): the Clusters of Orthologous Genes (COG) database, the Sanger Protein family (Pfam) database ((Punta et al., 2011)), and the Kyoto Encyclopedia of Genes and Genomes (KEGG) database (Kanehisa et al., 2014).

Dr. Lazar and I then sorted the genes of each annotated genome by function (e.g. stress response, viral genes, carbohydrate metabolism, etc.). Using published studies and database entries (COG, Pfam, and KEGG), I then constructed putative metabolic pathways for the MCG genomes.

RESULTS

A total of 150 genomes, bacterial and archaeal, were assembled from the samples taken. Using 16S rRNA genes, it was determined that of the genomes, 130 were bacterial and 20 were archaeal. Of the archaeal genomes, 6 were from the Miscellaneous Crenarchaeotal Group (MCG): two from MCG-1, one from MCG-6, two from MCG-7/17, and one from MCG-15. The six MCG genomes in this study were all more than complete (>100%) and ranged from 1.01 to 2.01 Mbp in size and had 1267 to 2649 genes, after correcting for completeness (Table 1).

Genome	Depth	Subgroup	Genome size (Mbp)	Coding regions
AD8-1	Shallow	MCG-6	1.54	1955
SG8-32-1	Shallow	MCG-1	2.01	2649
SG8-32-3	Shallow	MCG-1	1.39	1834
SMTZ-55	SMTZ	MCG-7/17	1.01	1278
SMTZ-80	SMTZ	MCG-7/17	1.15	1267
Deep-45	Deep	MCG-15	1.50	1781

Table 1. Genome summaries for MCG archaea in this study. Depth is indicated in the second column: shallow (8-12 cmbsf), SMTZ (sulfate to methane transition zone) (24-32 cmbsf), and deep (44-54 cmbsf). MCG subgroup classification is that proposed by Kubo et al. 2012.

All genes discussed below had sequence homologies above 20%, with some as high as 60%[3]. The results are summarized in Table II.

Nitrate and nitrite reduction

Each MCG from this study had at least one copy of *nirB*, with one of the MCG-6 genomes having 11 copies. The *nirB* gene encodes a single cytoplasmic protein that reduces nitrite (NO_2^-) to ammonium (NH_4^+) as the second part of a process called dissimilatory nitrate reduction to ammonium (DNRA) (Figure 2). The NirB polypeptide is occasionally found in complexes with NirD, whose function is unknown. Two copies of *nirD* were found in one of the MCG-6 genomes from the shallow sample.

The *nirB* gene is thought to be regulated by the gene *fnr*, which is a member of the Crp family of regulators (cAMP receptor protein). A total of 4 copies of this gene were found in the samples: one copy in an MCG-1 and three copies in an MCG-7/17.

[3] This is not atypical given the diversity of many enzymes. It is often the case that preservation of the active site and other binding sites is highest, while other segments experience lower selection pressure. Some archaeal enzymes have only 20% identities between phylotypes (Kappler, 2011). I have yet to analyze the active sites of the enzymes discussed here, but further studies will shed more light on the functions of these enzymes.

Another regulator of nirB is the 2-component system nar X/Q-narL/P, which was found in the other MCG-1 genome. Considered together, all of the genomes have *nirB* and half have purported *nirB* regulators.

A gene that was predominant in the upper layer and SMTZ, but absent in the deep sample, was *nosD*, a periplasmic nitrous oxide reductase accessory protein. The enzyme responsible for catalyzing the reduction of nitrous oxide (N_2O) to nitrogen gas (N_2) is NosZ, which was not detected in any genome (see discussion below). Nitrous oxide reductase genes (*nos* genes) were found in four of the six MCG genomes (Table 2 and Figure 2). The *nos* operon can be regulated by nsrR, DNR, or NnR; a total of six *nsrR* genes were found in three of the four genomes containing the *nosD* gene.

In addition to these major proteins, less abundant genes pertaining to inorganic metabolism of MCG were found as well. The Nar system of proteins is a group of membrane-bound nitrate reductases. The catalytic subunit of the nar operon in Archaea is NarG, which was found in one shallow MCG-1 genome. Other genes in the nar operon were found in the genomes, e.g. *narI*, *narL*, *narK*, and *narQ*- NarK/U is an N-oxyanion transporter. Periplasmic nitrate reductases are encoded by the gene *nap. Nap* genes were present in four of six genomes, though none had a copy of the gene thought to code for the catalytic subunit, *nap*A.

Sulfur metabolism

Although many enzymes involved in sulfur metabolism were detected, nearly all are involved in assimilatory processes (i.e. incorporation into biomass), which is not the

focus here. For example, copies of *aslA* were found in most of the genomes. This gene has be a serine protease and is likely not unique to any particular pathway (Szameit et al., 1999).

A gene of considerable interest that appeared in five of the six genomes is a gene only reliably identified by the COG ID 2041. Most databases indicate that this gene codes for a sulfite-oxidase gene. A gene often associated with COG 2041 was *yedY*. This is an enzyme that has been studied in pathogenic bacteria and has great sequence homology to other sulfite-oxidizing enzymes. However, when individual sequences were investigated, the highest sequence homology was with other archaea. For example, a sequence of 230 amino acids covered 82% of one of our sulfite-oxidase genes and they had a coverage of 82%, homology of 57%, and an e-value of 10^{-69} (accession no. YP_001403590). Comparison with other class 3 sulfite oxidases (see appendix) of cd_2109 (conserved domain) from *Sinorhizobium meliloti* gave much lower scores of 69% coverage, 28% homology, and an e-value of 10^{-20} (accession no. YP_007614251.1) (Finan et al, 2001).

		Copies						
		MCG subdivision						
		6	1	1	7/17	7/17	15	
	Gene	Shallow			SMTZ		Deep	Function
$NO_2^- \rightarrow NH_4^+$	*nirB/D*	2	13	5	3	3	1	cytoplasmic nitrite reductase
	fnr/CRP	0	0	1	0	3	0	potential nirB regulator
	narLQ	0	1	0	0	0	0	potential nirB regulator
$N_2O \rightarrow N_2$	*nosD/Y*	8	24	4	0	2	0	nitrous oxide reductase accessory proteins
	ygbA	1	1	4	0	0	0	putative nos-regulator/NO-stimulated promoter
$NO_3^- \rightarrow NO_2^-$	*napCDFH*	0	1	6	0	1	1	periplasmic nitrate reductase accessory proteins
	narGILQ	1	2	0	0	0	2	membrane-bound nitrate reductase subunits
Other	COG2041	1	5	5	0	3	1	potential sulfite oxidase
	narK/U	1	0	0	3	2	3	NO_3^-/NO_2^- transporter

Table 2. The major findings of this study. The left-most column indicates the processes in which the genes are involved. COG2041 is a sulfite-oxidase-like protein, closely related to YedY. NarQL is capable of regulating both *nirB* and *nap* operons and is categorized under both.

DISCUSSION

This study analyzed six complete MCG genomes from subsurface sediments taken from the WOR estuary. The large amount of genetic material provided us with a unique opportunity to investigate the metabolic capabilities of the MCG archaea with respect to nitrogen and sulfur cycling. Since sediments are anoxic, alternative electron acceptors must be used if the organism is going to generate ATP through respiration. No indications of the ability to use Fe(III) or Mn(IV) were detected in any genome. Additionally, the absence of *dsr* genes adds credence to the hypothesis that the MCG are not sulfate-reducers. This is consistent with Teske and Lloyd's hypothesis that MCG do

not change abundance with changing sulfate concentrations (K. G. Lloyd et al., 2011). Furthermore, no genes implicated in strictly aerobic processes were found, supporting the observations made by others that MCG may be strictly anaerobic archaea (Kubo et al., 2012).

No definite conclusions can be drawn about the potential for sulfite oxidation in the MCG. While a sulfite oxidizing enzyme (SOE)-derived enzyme is present in five of the six genomes (COG 2041), more evidence is necessary before a conclusion can be reached concerning the MCG's ability to oxidize sulfite. Recent studies have shown that class I SOEs have a subdivision that is encoded by the gene *yedY.* This enzyme, although having the structure of a SOE, does not oxidize sulfite in a laboratory setting, though it may have the ability to reduce dimethylsulfoxide (Kappler, 2011). However, the archaeal SOEs that have been examined thus far (class III, cd_2109) do not seem to have this odd characteristic (Wilson & Kappler, 2009), suggesting that this gene may be functional in the MCG. Sulfite oxidation would be a potentially beneficial in light of a recent study by Lloyd that shows MCG can degrade amino acids, a process which is known to release sulfites, which are toxic to cells. (Dahl & Friedrich, 2007; Lloyd et al., 2013).

The biologically mediated nitrogen cycle is controlled by a large number of genes, most of which are shown in Figure 2. Some of these processes are exclusively carried out in bacteria and archaea (e.g. nitrogen fixation), while some occur in all three domains (e.g. nitrate reduction). To what extent the archaea play a role in the nitrogen cycle is unknown, but evidence suggests they are capable of carrying out all reactions in Figure 2, as are the bacteria. (Cabello, Roldán, & Moreno-Vivián, 2004) Given their

abundance, the MCG analyzed here seem to be capable of significantly affecting the global nitrogen cycle.

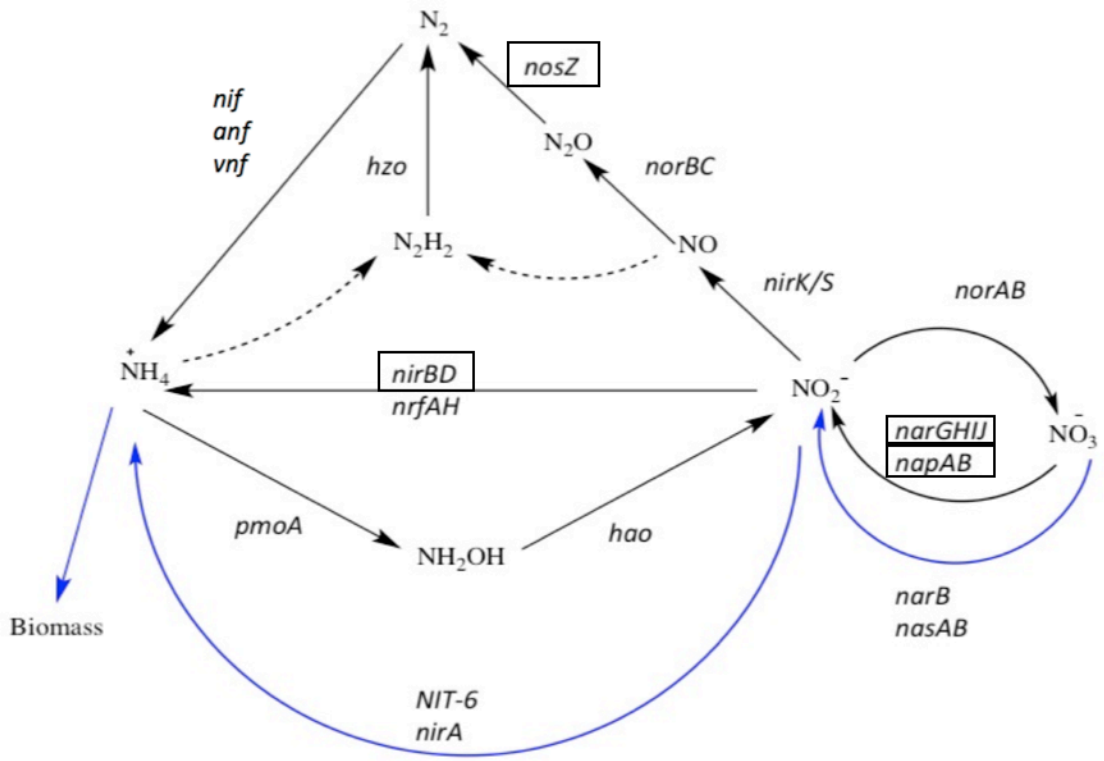


Figure 2 The biologically mediated nitrogen cycle. The genes known to catalyze a given reaction are shown between their reactant and product. Processes in blue are assimilatory (converted to biomass) while processes in black are dissimilatory (not converted to biomass). Genes found in the MCG genomes in this study are boxed.

The MCG have the ability to reduce nitrite to ammonium

The most supported finding from this study is that the MCG of the subdivisions 1, 6, 7/17, and 15 are all capable of reducing nitrite to ammonium in a dissimilatory manner. This is part of a process called dissimilatory reduction of nitrate to ammonium (DNRA) or ammonification. The process is often referred to as a single reaction, but it is actually carried out by two sets of enzymes. NO_3^- is first reduced to NO_2^- by either Nap or Nar

reductases (see below; Figure 2). Then, either NirB or NrfA reduces NO_2^- directly to NH_4^+. Copies of *nirB*, which is a single polypeptide, were found in every genome in this study: this is strong evidence that these MCG can reduce nitrite to ammonium.

NirB contains the prosthetic group siroheme (a modified form of heme) which is found mostly in sulfite and nitrite reductases (Murphy, Siegel, Tove, & Kamin, 1974). An important distinction must be made between NirB and NrfA (not present in any genome): while they both carry out the same reaction, they do so by entirely different mechanisms (Mohan & Cole, 2007). Much more is known about NrfA than NirB and many questions relevant to this study are currently unanswered. For example, NrfA has been reported to produce NO and N_2O as byproducts of ammonification (Corker & Poole, 2003), although studies countering this have also been done (Gilberthorpe & Poole, 2008). If NirB also produces N_2O as a byproduct, this could explain, in part, the presence of the *nos* genes, though this side product may only account for approximately 1% of NO_2^- reduced by nitrite reductases at ecologically relevant pHs (Cole, 1990; Rütting, Boeckx, Müller, & Klemedtsson, 2011)

Nitrous oxide reductase

Of the enzymes known to take part in denitrification (the stepwise reduction of NO_3^- to N_2 by use as a terminal electron acceptor), Nos has been shown to be the least common, or at least the least detected in environmental samples. (Bru et al., 2011; Henry, Bru, Stres, Hallet, & Philippot, 2006) However, numerous copies of the *nosZ* accessory protein, *nosD*, were found in all of the shallow genomes, and *nosY* was seen in one of the

MCG-1 genomes. The absence of the catalytic subunit, nosZ, is not particularly surprising given recent studies involving this gene. It has been shown that *nosZ* has a tremendous amount of sequence variability and that the original PCR primers used to assay for *nosZ* presence fail to detect the gene in some organisms that are known to reduce N_2O to N_2. (Jones, Graf, Bru, Philippot, & Hallin, 2013; Sanford et al., 2012; Scala & Kerkhof, 1999) Furthermore, this same scenario has been reported in other archaeal genomes: *nosDY* is present in *Haloferax volcanii*, a euryarchaeote; but *nosZ* is absent, although it is not known if *H. volcanii* can use N_2O as an electron acceptor (Zumft and Körner, 2007).

It therefore seems entirely possible that the genomes coding for the nosZ accessory proteins nosY and nosD, especially multiple copies, contain some type of nitrous oxide reductase not recognized during genome annotation. This hypothesis is further bolstered by a discovery from 1993 in which a novel type of nitrous oxide reductase was found in the Crenarchaeote *Pyrobacculum aerophilum* (Völkl et al., 1993).

The potential presence of only part of the denitrification pathway is not uncommon in prokaryotes; in fact, partial denitrifiers are more common than complete dentirifiers since each step generates a proton motive force that is used in ATP synthesis (Shapleigh, 2013). The reduction of N_2O to N_2 is an exergonic and energetically favorable reaction (ΔG_0'= -339.5 kJ/mol) and therefore does not need to be coupled to another reaction to be biologically useful (Pauleta, Dell'Acqua, & Moura, 2013).

Nitrate reductase genes are largely absent

There are two enzymes known to reduce nitrate to nitrite in a dissimilatory fashion: NarGHI, which is membrane bound, and NapAB, which is periplasmic. Interestingly, the Nar complex has a reversed orientation in archaea, with the active site and catalytic subunits (G and H) on the outside of the cytoplasm (Richardson et al., 2007). The two subunits coded for by *narG* and *narH* form a dimer, which binds to the membrane-bound NarI. An MCG-1 from this study had a copy of the catalytic subunit gene, *narG* and the MCG-15 had two copies of *narI*. Genes coding for the subunit NarH were entirely absent from all genomes. NarQL is capable of regulating both the *nir* and *nap* operons. (Philippot, 2002; Sparacino-Watkins, Stolz, & Basu, 2013). NarQL is a two-component regulatory system, with the sensor, NarQ, embedded in the membrane and NarL binding to DNA in the presence of nitrate or nitrite. Interestingly, the gene *narQ* was detected in the MCG-6 and an MCG-1, and *narL* was present in the same MCG-1 genome.

The Nap nitrate reductase is typically comprised of two subunits (A and B), however, functioning NapA, without NapB, has been found in some species of bacteria ((Dias et al., 1999). The only *nap* genes detected in our archaeal genomes were those that code for accessory proteins (C,D,F,H). NapC and H, both of which are membrane cytochromes, were the most abundant (accounting for seven of the nine copies), while one copy of both napD and F were found; both of which are involved in the maturation of NapA (Mohan & Cole, 2007).

In these MCG genomes, the evidence for NO_3^- reduction to NO_2^- is therefore not particularly compelling. The almost uniform absence of catalytic subunits for both periplasmic and membrane nitrate reductases implies that this is not a reaction the MCG

subgroups analyzed here can carry out using known enzymes. Though it is not possible to say for certain why genes coding for accessory proteins are present in some genomes, even though the genes coding for catalytic subunits are absent, there are potential explanations. The MCG may have formerly been capable of reducing nitrate to nitrite and have since lost some of the genes necessary to carry out the reaction. It is also possible that the catalytic subunits are of a novel type and are therefore not recognized in any known database. Overall, more evidence is needed to determine whether or not the MCG can reduce nitrate to nitrite.

Transport proteins

In terms of inorganic ion transport, four of the six genomes had at least one copy of the *narK/U* gene. This gene family codes for two different types of nitrate transporter: one variant is a NO_3^-/NO_2^- antiporter; the other a NO_3^-/H^+ symporter (Wood, Alizadeh, Richardson, Ferguson, & Moir, 2002). Though referred to as nitrate transporters, evidence suggests that other N-oxyanions (e.g. nitrite) may also be transported (Richardson, van Spanning, & Ferguson, 2007).

Based on the absence on NO_2^- transporters in nitrite reducing bacteria, it has been proposed that nitrite can freely diffuse through membranes in their protonated form, as HNO_2, or that nitrite does not require a specific transporter (Moir & Wood, 2001). We therefore conclude that nitrite has some mechanism of getting into the cell and the absence of known transporters likely does not affect the MCG's ability to reduce nitrite.

The interaction between genes and environment

The genes found in this study support the hypothesis that some MCG archaea have the ability to reduce nitrite to ammonium (nitrite ammonification) and respire nitrous oxide via reduction to N_2. Though involving similar compounds, these two processes are not physiologically connected to one another. Notably, this exact combination of processes (the last step of denitrification and nitrite ammonification) is found in *Wolinella succinogenes* (Simon, Einsle, Kroneck, & Zumft, 2004). Ecologically, these two processes have opposite consequences: the production of ammonium keeps nitrogen in the local environment, since highly reduced nitrogen compounds are the only ones capable of integration to biomass (Cabello et al., 2004). On the other hand, the production of N_2 leads to a loss of nitrogen from an ecosystem, since N_2 is an inert gas and can only be used by diazotrophs (Mohan & Cole, 2007). Whether this is significant (i.e. whether fixed nitrogen is scarce) depends on the ecosystem; however, it is doubtful that the preservation of fixed nitrogen is of great importance in the WOR estuary, since estuaries are nutrient rich and have high nitrogen concentrations (Sørensen, 1978).

The WOR estuary is rich in organic carbon (Supplementary Figure 1), and since organic rich-sediments are typically rich in NH_4^+, nitrite assimilation is expected to be minimal (Jørgensen, 1989). This means that nitrite can be reduced by denitrification, anammox (anaerobic oxidation of ammonium) or DNRA. Based on the abundance of *nirB* and the absence of *hzo*, *norBC,* and *nirK/S*, it seems that DNRA is the preferred method of NO_2^- reduction for MCG. Given that marine sediments are have low reduction potentials (i.e. are reducing (Sørensen, 1978)) DNRA is a more favorable reaction from

an energetics point of view. The siroheme prosthetic group of *nirB* transfers six electrons to each NO_2^-, while denitrification only transfers two or three. Since electrons are abundant, it makes sense to use reactions that involve more electron transfers to each electron acceptor (NO_2^-) (Mohan & Cole, 2007). NirB often uses NADH as the electron donor, therein regenerating NAD^+, a necessary oxidant for metabolic processes. At a molecular level, this is very different than denitrification: denitrification is respiratory, in that it produces a proton motive force; DNRA on the other hand, does not, instead generating less ATP by substrate-level phosphorylation; it is thus more accurate to call it "dissimilatory" (Schmidt & Schaechter, 2012). Overall, this reaction serves as an "electron sink" and helps maintain the cellular redox balance (Simon, 2002; Zumft, 1997).

Furthermore, estuarine water columns and sediments have higher concentrations of NO_3^- than dissolved O_2, which allows denitrification to account for more carbon oxidation than aerobic processes (Jorgensen & Sorensen, 1985). It therefore seems reasonable to assume a large presence of NO_3^- reducing microbes. It could be the case that these denitrifying bacteria and archaea can only carry out the first few steps of denitrification (nitrate, nitrite, and nitric oxide reduction). The release of N_2O by these partial denitrifiers could provide the terminal electron acceptor for these MCG.

Knowledge about archaeal and bacterial metabolisms may allow us to exploit these reactions to try and reverse some of the damage our species has done to our planet (Atlas, 1995; Finneran, Housewright, & Lovley, 2002). Nitrate runoff is a very common problem associated with some fertilization methods. The excess nitrate in ecosystems enriches for denitrifiers, which leads to a tremendous increase in N_2O emissions, a very

potent greenhouse gas, nearly 300 times as harmful than CO_2 (Naqvi et al., 2000). Knowing that some organisms, like the MCG in this study, can potentially respire solely on N_2O may allow us to develop methods to reduce natural and anthropogenic N_2O emissions.

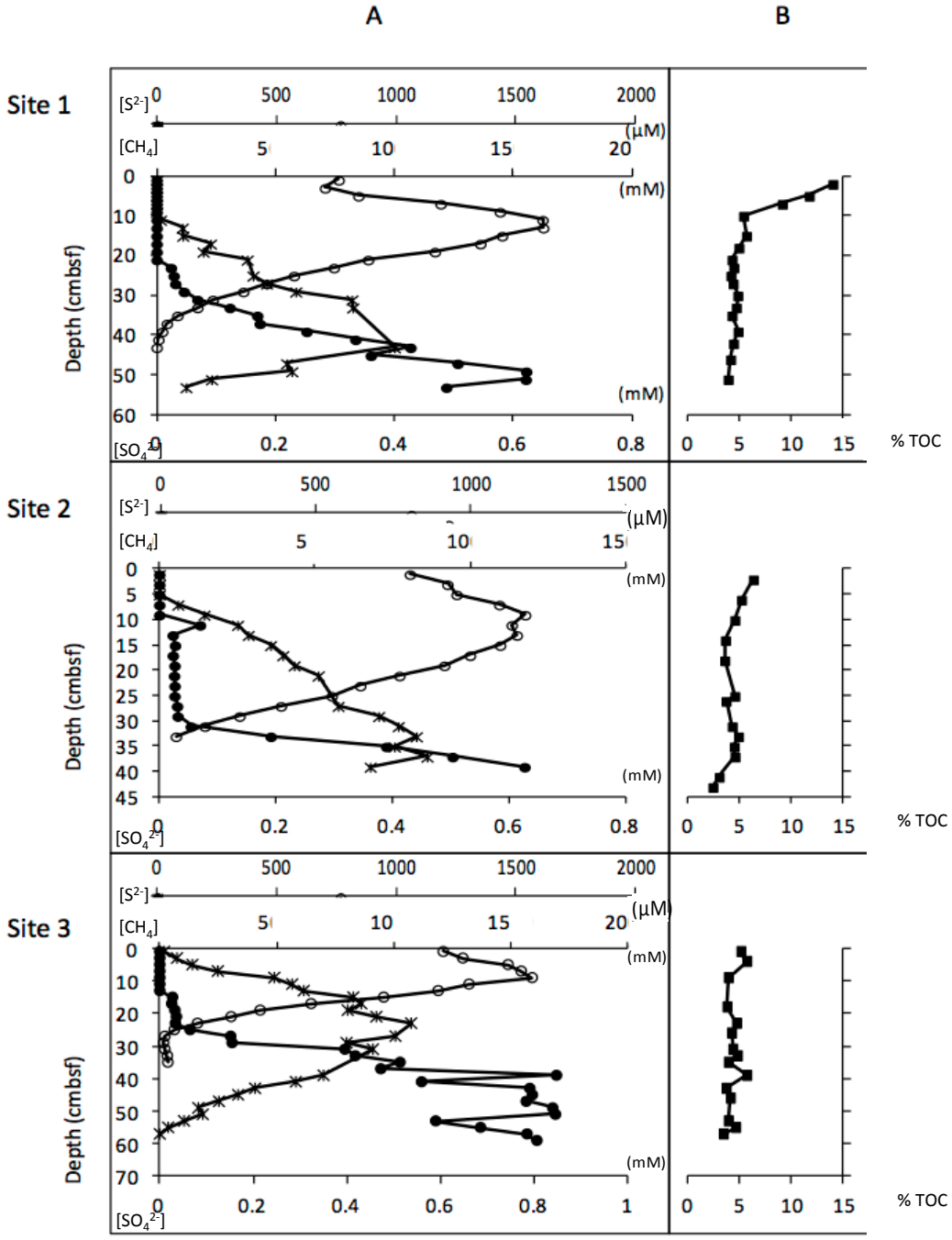
This work presents results that suggest for the first time the capacity for MCG involvement in nitrogen cycling. Though our genomic evidence implicates MCG in DNRA and N_2O reduction, these results cannot yet be generalized to all MCG, since so little is known about this very diverse group of archaea. Despite recent attempts (Gagen et al., 2013), the MCG have yet to be cultured; and while metagenomics is a remarkably powerful tool for analyzing uncultured organisms and whole communities, there are significant limitations that must be acknowledged before too much weight is given to any finding. Metagenomics only tells us what genes are present, not whether or not they are being used; a metatranscriptomics-based study of the MCG would help address this.

Often, genomics approaches (i.e. metagenomics, metatranscriptomics, and metaproteomics) are deemed to be "culture-independent" techniques. This is true in that they do not require the organism to be cultured before it is analyzed. However, it is also slightly misleading because these techniques do still rely on cultured organisms: the information we have about microbial genes and proteins that we use for annotation come from cultured organisms. The findings from metagenomic studies could potentially aid in culturing attempts, which would improve these databases; this limitation should be kept in mind when looking at archaeal genomes. Furthermore, half of each genome we examined was comprised of genes with unknown functions- what could these genes be

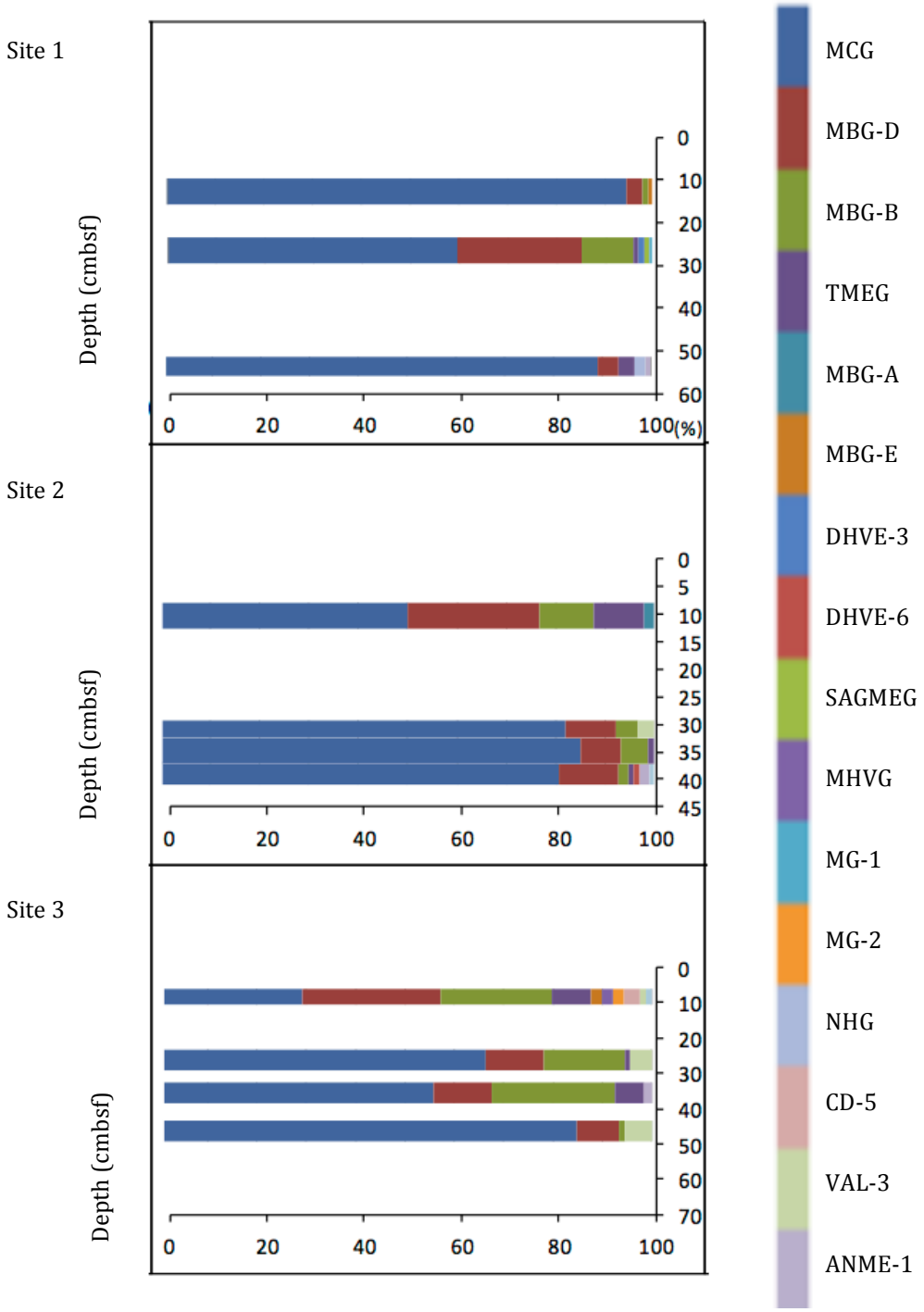
doing? In this way, metagenomics is a limited approach, perhaps underestimating the capabilities of these microbes.

Despite all of the above caveats, metagenomics is still a remarkably powerful approach, allowing us to probe parts of nature that have previously been off limits. The evidence we have analyzed using this technique has shown that the MCG archaea likely take part in several parts of the global nitrogen cycle. Given their abundance in marine sediments, these findings may allow for us to refine current geochemical models so that we may better understand the role archaea play in the cycling of elements essential for life.

SUPPLEMENTARY INFORMATION



Supplementary Figure 1. Geochemical measurements of sample sites at White Oak River (WOR) estuary. (A) shows concentrations of CH_4 (•, mM); SO_4^{2-} (O, mM); and S^{2-} (★, µM). (B) shows percentage of total organic carbon (TOC) as a function of depth. Figure is adapted from Lazar et al., in press.



Supplementary Figure II. Archaeal phylotype abundance at three sampling sites in WOR estuary. MCG: Miscellaneous Crenarchaeotal Group; MBG: Marine Benthic Group; TMEG: Terrestrial Misc. Euryarchaeotal Group; DHVE: Deep-sea Hydrothermal Vent Euryarchaeotal Group; SAGMEG: South African Goldmine Euryarchaeotal Group; MHVG: Marine Hydrothermal Vent Group; MG: Marine Group; NHG: Novel *Halobacteriales* Group; CD: Candidate Division; VAL: Valkea Kotinen; ANME: Anaerobic Methanotrophic archaea.

REFERENCES

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology*, *215*(3), 403–410. doi:10.1016/S0022-2836(05)80360-2

Atlas, R. M. (1995). Bioremediation of petroleum pollutants. *International Biodeterioration & Biodegradation*, *35*(1–3), 317–327. doi:10.1016/0964-8305(95)00030-9

Barns, S. M., Delwiche, C. F., Palmer, J. D., & Pace, N. R. (1996). Perspectives on archaeal diversity, thermophily and monophyly from environmental rRNA sequences. *Proceedings of the National Academy of Sciences*, *93*(17), 9188–9193.

Bennett, S. (2004). Solexa Ltd. *Pharmacogenomics*, *5*(4), 433–438. doi:10.1517/14622416.5.4.433

Biddle, J. F., Lipp, J. S., Lever, M. A., Lloyd, K. G., Sørensen, K. B., Anderson, R., … Hinrichs, K.-U. (2006). Heterotrophic Archaea dominate sedimentary subsurface ecosystems off Peru. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(10), 3846–3851. doi:10.1073/pnas.0600035103

Bru, D., Ramette, A., Saby, N. P. A., Dequiedt, S., Ranjard, L., Jolivet, C., … Philippot, L. (2011). Determinants of the distribution of nitrogen-cycling microbial communities at the landscape scale. *Isme Journal*, *5*(3), 532–542. doi:10.1038/ismej.2010.130

Cabello, P., Roldán, M. D., & Moreno-Vivián, C. (2004). Nitrate reduction and the nitrogen cycle in archaea. *Microbiology*, *150*(11), 3527–3546. doi:10.1099/mic.0.27303-0

Cole, J. A. (1990). Physiology, Biochemistry and Genetics of Nitrate Dissimilation to Ammonia. In N. P. Revsbech & J. Sørensen (Eds.), *Denitrification in Soil and Sediment* (pp. 57–76). Springer US. Retrieved from http://link.springer.com/chapter/10.1007/978-1-4757-9969-9_4

Corker, H., & Poole, R. K. (2003). Nitric Oxide Formation by Escherichia coli DEPENDENCE ON NITRITE REDUCTASE, THE NO-SENSING REGULATOR Fnr, AND FLAVOHEMOGLOBIN Hmp. *Journal of Biological Chemistry*, *278*(34), 31584–31592. doi:10.1074/jbc.M303282200

Dahl, C., & Friedrich, C. G. (2007). *Microbial Sulfur Metabolism*. Springer.

Delong, E. F. (1994). High abundance of archaea in Antarctic marine picoplankton. *Nature*, *371*, 695–697.

Dias, J. M., Than, M. E., Humm, A., Huber, R., Bourenkov, G. P., Bartunik, H. D., … Romão, M. J. (1999). Crystal structure of the first dissimilatory nitrate reductase at 1.9 Å solved by MAD methods. *Structure*, *7*(1), 65–79. doi:10.1016/S0969-2126(99)80010-0

Dong, Y., Kumar, C. G., Chia, N., Kim, P.-J., Miller, P. A., Price, N. D., … Fouke, B. W. (2014). Halomonas sulfidaeris-dominated microbial community inhabits a 1.8 km-deep subsurface Cambrian Sandstone reservoir. *Environmental Microbiology*, n/a–n/a. doi:10.1111/1462-2920.12325

Edgcomb, V. P., Beaudoin, D., Gast, R., Biddle, J. F., & Teske, A. (2011). Marine subsurface eukaryotes: the fungal majority. *Environmental Microbiology*, *13*(1), 172–183. doi:10.1111/j.1462-2920.2010.02318.x

Elshahed, M. S., Senko, J. M., Najar, F. Z., Kenton, S. M., Roe, B. A., Dewers, T. A., … Krumholz, L. R. (2003). Bacterial Diversity and Sulfur Cycling in a Mesophilic Sulfide-Rich Spring. *Applied and Environmental Microbiology*, *69*(9), 5609–5621. doi:10.1128/AEM.69.9.5609-5621.2003

Finneran, K. T., Housewright, M. E., & Lovley, D. R. (2002). Multiple influences of nitrate on uranium solubility during bioremediation of uranium-contaminated subsurface sediments. *Environmental Microbiology*, *4*(9), 510–516. doi:10.1046/j.1462-2920.2002.00317.x

Gagen, E. J., Huber, H., Meador, T., Hinrichs, K.-U., & Thomm, M. (2013). Novel Cultivation-Based Approach To Understanding the Miscellaneous Crenarchaeotic Group (MCG) Archaea from Sedimentary Ecosystems. *Applied and Environmental Microbiology*, *79*(20), 6400–6406. doi:10.1128/AEM.02153-13

Gilberthorpe, N. J., & Poole, R. K. (2008). Nitric Oxide Homeostasis in Salmonella typhimurium ROLES OF RESPIRATORY NITRATE REDUCTASE AND FLAVOHEMOGLOBIN. *Journal of Biological Chemistry*, *283*(17), 11146–11154. doi:10.1074/jbc.M708019200

Henry, S., Bru, D., Stres, B., Hallet, S., & Philippot, L. (2006). Quantitative Detection of the nosZ Gene, Encoding Nitrous Oxide Reductase, and Comparison of the

Abundances of 16S rRNA, narG, nirK, and nosZ Genes in Soils. *Applied and Environmental Microbiology*, *72*(8), 5181–5189. doi:10.1128/AEM.00231-06

Hershberger, K. L., Barns, S. M., Reysenbach, A.-L., Dawson, S. C., & Pace, N. R. (1996). Wide diversity of Crenarchaeota. *Nature*, *384*(6608), 420–420. doi:10.1038/384420a0

Jones, C. M., Graf, D. R., Bru, D., Philippot, L., & Hallin, S. (2013). The unaccounted yet abundant nitrous oxide-reducing microbial community: a potential nitrous oxide sink. *The ISME Journal*, *7*(2), 417–426. doi:10.1038/ismej.2012.125

Jorgensen, B., & Sorensen, J. (1985). Seasonal Cycles of O-2, No-3-3 and so-4(2-) Reduction in Estuarine Sediments - the Significance of an No-3-Reduction Maximum in Spring. *Marine Ecology Progress Series*, *24*(1-2), 65–74. doi:10.3354/meps024065

Jørgensen, K. S. (1989). Annual Pattern of Denitrification and Nitrate Ammonification in Estuarine Sediment. *Applied and Environmental Microbiology*, *55*(7), 1841–1847.

Kanehisa, M., Goto, S., Sato, Y., Kawashima, M., Furumichi, M., & Tanabe, M. (2014). Data, information, knowledge and principle: back to metabolism in KEGG. *Nucleic Acids Research*, *42*(Database issue), D199–205. doi:10.1093/nar/gkt1076

Kappler, U. (2011). Bacterial sulfite-oxidizing enzymes. *Biochimica et Biophysica Acta (BBA) - Bioenergetics*, *1807*(1), 1–10. doi:10.1016/j.bbabio.2010.09.004

Kubo, K., Lloyd, K. G., Biddle, J. F., Amann, R., Teske, A., & Knittel, K. (2012). Archaea of the Miscellaneous Crenarchaeotal Group are abundant, diverse and

widespread in marine sediments. *The ISME Journal*, *6*(10), 1949–1965. doi:10.1038/ismej.2012.37

Lipp, J. S., Morono, Y., Inagaki, F., & Hinrichs, K.-U. (2008). Significant contribution of Archaea to extant biomass in marine subsurface sediments. *Nature*, *454*(7207), 991–994. doi:10.1038/nature07174

Lloyd, K. G., Alperin, M. J., & Teske, A. (2011). Environmental evidence for net methane production and oxidation in putative ANaerobic MEthanotrophic (ANME) archaea. *Environmental Microbiology*, *13*(9), 2548–2564. doi:10.1111/j.1462-2920.2011.02526.x

Lloyd, K. G., Biddle, J., & Teske, A. (2011). Environmental distribution, abundance and activity of the Miscellaneous Crenarchaeotal Group. *AGU Fall Meeting Abstracts*, *-1*, 0398.

Lloyd, K. G., Schreiber, L., Petersen, D. G., Kjeldsen, K. U., Lever, M. A., Steen, A. D., … Jørgensen, B. B. (2013). Predominant archaea in marine sediments degrade detrital proteins. *Nature*, *496*(7444), 215–218. doi:10.1038/nature12033

Miller, T. L., Wolin, M. J., Conway de Macario, E., & Macario, A. J. (1982). Isolation of Methanobrevibacter smithii from human feces. *Applied and Environmental Microbiology*, *43*(1), 227–232.

Moir*, J. W. B., & Wood, N. J. (2001). Nitrate and nitrite transport in bacteria. *Cellular and Molecular Life Sciences CMLS*, *58*(2), 215–224. doi:10.1007/PL00000849

Murphy, M. J., Siegel, L. M., Tove, S. R., & Kamin, H. (1974). Siroheme: A New Prosthetic Group Participating in Six-Electron Reduction Reactions Catalyzed

by Both Sulfite and Nitrite Reductases. *Proceedings of the National Academy of Sciences*, *71*(3), 612–616.

Naqvi, S. W. A., Jayakumar, D. A., Narvekar, P. V., Naik, H., Sarma, V. V. S. S., D'Souza, W., … George, M. D. (2000). Increased marine production of N2O due to intensifying anoxia on the Indian continental shelf. *Nature*, *408*(6810), 346–349. doi:10.1038/35042551

Nercessian, O., Bienvenu, N., Moreira, D., Prieur, D., & Jeanthon, C. (2005). Diversity of functional genes of methanogens, methanotrophs and sulfate reducers in deep-sea hydrothermal environments. *Environmental Microbiology*, *7*(1), 118–132. doi:10.1111/j.1462-2920.2004.00672.x

Parkes, R. J., Webster, G., Cragg, B. A., Weightman, A. J., Newberry, C. J., Ferdelman, T. G., … Fry, J. C. (2005). Deep sub-seafloor prokaryotes stimulated at interfaces over geological time. *Nature*, *436*(7049), 390–394. doi:10.1038/nature03796

Pauleta, S. R., Dell'Acqua, S., & Moura, I. (2013). Nitrous oxide reductase. *Coordination Chemistry Reviews*, *257*(2), 332–349. doi:10.1016/j.ccr.2012.05.026

Philippot, L. (2002). Denitrifying genes in bacterial and Archaeal genomes. *Biochimica et Biophysica Acta*, *1577*(3), 355–376.

Punta, M., Coggill, P. C., Eberhardt, R. Y., Mistry, J., Tate, J., Boursnell, C., … Finn, R. D. (2011). The Pfam protein families database. *Nucleic Acids Research*, *40*(D1), D290–D301. doi:10.1093/nar/gkr1065

Rütting, T., Boeckx, P., Müller, C., & Klemedtsson, L. (2011). Assessment of the importance of dissimilatory nitrate reduction to ammonium for the

terrestrial nitrogen cycle. *Biogeosciences*, *8*(7), 1779–1791. doi:10.5194/bg-8-1779-2011

Sanford, R. A., Wagner, D. D., Wu, Q., Chee-Sanford, J. C., Thomas, S. H., Cruz-Garcia, C., … Loffler, F. E. (2012). Unexpected nondenitrifier nitrous oxide reductase gene diversity and abundance in soils. *Proceedings of the National Academy of Sciences*, *109*(48), 19709–19714. doi:10.1073/pnas.1211238109

Scala, D. J., & Kerkhof, L. J. (1999). Diversity of Nitrous Oxide Reductase (nosZ) Genes in Continental Shelf Sediments. *Applied and Environmental Microbiology*, *65*(4), 1681–1687.

Schmidt, T. M., & Schaechter, M. (2012). *Topics in Ecological and Environmental Microbiology*. Academic Press.

Shapleigh, D. J. P. (2013). Denitrifying Prokaryotes. In E. Rosenberg, E. F. DeLong, S. Lory, E. Stackebrandt, & F. Thompson (Eds.), *The Prokaryotes* (pp. 405–425). Springer Berlin Heidelberg. Retrieved from http://link.springer.com/referenceworkentry/10.1007/978-3-642-30141-4_71

Simon, J. (2002). Enzymology and bioenergetics of respiratory nitrite ammonification. *FEMS Microbiology Reviews*, *26*(3), 285–309.

Simon, J., Einsle, O., Kroneck, P. M. H., & Zumft, W. G. (2004). The unprecedented nos gene cluster of Wolinella succinogenes encodes a novel respiratory electron transfer pathway to cytochrome c nitrous oxide reductase. *FEBS Letters*, *569*(1–3), 7–12. doi:10.1016/j.febslet.2004.05.060

Sørensen, J. (1978). Capacity for Denitrification and Reduction of Nitrate to Ammonia in a Coastal Marine Sediment. *Applied and Environmental Microbiology*, *35*(2), 301–305.

Sparacino-Watkins, C., Stolz, J. F., & Basu, P. (2013). Nitrate and periplasmic nitrate reductases. *Chemical Society Reviews*, *43*(2), 676–706. doi:10.1039/C3CS60249D

Strous, M., Fuerst, J. A., Kramer, E. H., Logemann, S., Muyzer, G., van de Pas-Schoonen, K. T., … Jetten, M. S. (1999). Missing lithotroph identified as new planctomycete. *Nature*, *400*(6743), 446–449. doi:10.1038/22749

Szameit, C., Miech, C., Balleininger, M., Schmidt, B., Figura, K. von, & Dierks, T. (1999). The Iron Sulfur Protein AtsB Is Required for Posttranslational Formation of Formylglycine in the Klebsiella Sulfatase. *Journal of Biological Chemistry*, *274*(22), 15375–15381.

Teske, A., & Sørensen, K. B. (2007). Uncultured archaea in deep marine subsurface sediments: have we caught them all? *The ISME Journal*, *2*(1), 3–18. doi:10.1038/ismej.2007.90

Völkl, P., Huber, R., Drobner, E., Rachel, R., Burggraf, S., Trincone, A., & Stetter, K. O. (1993). Pyrobaculum aerophilum sp. nov., a novel nitrate-reducing hyperthermophilic archaeum. *Applied and Environmental Microbiology*, *59*(9), 2918–2926.

Whitman, W. B., Coleman, D. C., & Wiebe, W. J. (1998). Prokaryotes: The unseen majority. *Proceedings of the National Academy of Sciences*, *95*(12), 6578–6583.

Wilson, J. J., & Kappler, U. (2009). Sulfite oxidation in Sinorhizobium meliloti. *Biochimica et Biophysica Acta (BBA) - Bioenergetics*, *1787*(12), 1516–1525. doi:10.1016/j.bbabio.2009.07.005

Woese, C. R., & Fox, G. E. (1977). Phylogenetic structure of the prokaryotic domain: The primary kingdoms. *Proceedings of the National Academy of Sciences*, *74*(11), 5088–5090. doi:10.1073/pnas.74.11.5088

Woese, C. R., Kandler, O., & Wheelis, M. L. (1990). Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. *Proceedings of the National Academy of Sciences*, *87*(12), 4576–4579. doi:10.1073/pnas.87.12.4576

Wood, N. J., Alizadeh, T., Richardson, D. J., Ferguson, S. J., & Moir, J. W. B. (2002). Two domains of a dual-function NarK protein are required for nitrate uptake, the first step of denitrification in Paracoccus pantotrophus. *Molecular Microbiology*, *44*(1), 157–170. doi:10.1046/j.1365-2958.2002.02859.x

Zumft, W. G. (1997). Cell biology and molecular basis of denitrification. *Microbiology and Molecular Biology Reviews*, *61*(4), 533–616.